

**ELECTRONIC FILING**

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

|                  |   |                      |                               |
|------------------|---|----------------------|-------------------------------|
| Application of:  | Barbara Ensoli  | Confirmation No.:    | 9400                          |
| Application No.: | 09/555,534  | Art Unit:            | 1648                          |
| Filed:           | May 31, 2000  | Examiner:            | Humphrey, Louise Wang Zhiying |
| For:             | HIV TAT, OR DERIVATIVES<br>THEREOF FOR PROPHYLACTIC<br>AND THERAPEUTIC<br>VACCINATION | Attorney Docket No.: | 11340-003-999                 |

**FOURTH DECLARATION OF MAURO MAGNANI, Ph.D.  
UNDER 37 C.F.R. § 1.132**

---

Mail Stop AF  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

Sir:

I, MAURO MAGNANI, Ph.D., do declare as follows:

1. I am currently Professor of Biochemistry, Director of the section Biotechnology, Vice-Rector of the University of Urbino, Urbino, Italy, from 2001 to September 24, 2009. I have over thirty years of experience as a biochemist in the research and development of products and applications useful in the biotechnology and pharmaceutical industries. I am included in the official list of professional biologists in Italy with n. 017484 "Ordine Nazionale Biologi," and I am a Technical Director nominated by the "Agenzia Italiana del Farmaco, AIFA" with n. AIDT-19/2005. My education and experience are summarized on my Curriculum Vitae, which is attached hereto as Exhibit 1.

2. I have collaborated with Dr. Barbara Ensoli, who is the inventor of the above-identified application No. 09/555,534 (hereinafter "the '534 application"), in the development of methods to produce recombinant, biologically active HIV Tat protein. I also supervise and have supervised the production of such biologically active HIV Tat protein, according to good manufacturing practices (GMP), for use in human clinical trials. In particular, I have

experience in reversed-phase high pressure liquid chromatography (RP-HPLC) purification of biologically active HIV Tat protein, which my laboratory has carried out many times.

3. This Fourth Declaration is intended to supplement the Third Declaration of Mauro Magnani, Ph.D. Under 37 C.F.R. § 1.132 ("Third Magnani Declaration") filed January 8, 2009 for the '534 application. In particular, this Fourth Declaration provides further evidence that one skilled in the art as of December 1, 1997 would understand that an acid exchange reaction such as that described in U.S. Patent No. 5,646,120 by Sumner-Smith *et al.* ("Sumner-Smith *et al.*") would be expected to destroy the biological activity of an HIV Tat protein purified by RP-HPLC, even if the acid exchange reaction is performed before a lyophilization and resuspension step such as that described in Chang *et al.*, "HIV-1 Tat protein exits from cells via a leaderless secretory pathway and binds to extracellular matrix-associated heparin sulphate proteoglycans through its basic region," AIDS, 1997, 11(12):1421-1431 ("Chang *et al.*"). Indeed, as discussed below, it is the performance of the acid exchange reaction and consequent change of pH in an aqueous environment that promotes loss of biological activity of the HIV Tat protein, regardless of whether the acid exchange reaction occurs just after RP-HPLC or just after lyophilization and resuspension.

4. As disclosed in the specification of the '534 application (see page 24, lines 21-25), the conformation and biological activity of an HIV Tat protein is strongly affected by the protein's cysteine rich region. When oxidized, the cysteine residues of the HIV Tat protein form intra- and inter-molecular disulfide bonds, thus modifying the conformation of the HIV Tat protein, causing aggregation and loss of biological activity. At an acidic pH, the HIV Tat protein is reduced and in a random coil conformation; as the pH increases, the protein tends to change its conformation, forming a non-biologically active structure stabilized by disulfide bonds.

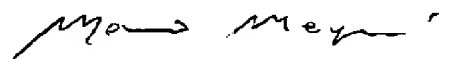
5. Sumner-Smith *et al.* discloses treating a column-eluted oligopeptide to exchange the cleavage acid, such as trifluoroacetic acid (TFA), with a pharmaceutically acceptable acid, such as acetic acid, to provide a water soluble salt of the peptide (see col. 9, lines 57-61). Such an acid exchange reaction, exchanging TFA for an acid like acetic acid, I understand to be a common practice for peptides, but not for proteins. The reason for this is that in general, proteins must retain their three-dimensional conformation in order to retain their biological activity, whereas short peptides often are not reliant on their three-

dimensional conformation for their biological activity or can recover their biological activity more easily than larger proteins. Acetic acid, which has a pKa of 4.76, is much less acidic than TFA, which has a pKa of 0.3. A person skilled in the art would understand that even a small increase in pH could cause a significant change in a protein's hydrogen exchange rate with water and changes in the protein conformation and biological activity. Indeed, a simple change in pH from 3.3 to 5.8 causes an increase in the hydrogen exchange rate with water on the order of 250-fold for an HIV Tat protein, as determined by NMR spectroscopy, with consequent appearance of local conformation (see Shojania *et al.*, "HIV-1 Tat is a natively unfolded protein," J Biol. Chem., 2006, 281(13):8347-8356, at the paragraph spanning pages 8354-8355, copy of which is attached as Exhibit 2). One would expect that, under acid exchange conditions, the presence of a reducing agent would be necessary, although not necessarily sufficient, for the HIV Tat protein to maintain a non-oxidized, non-aggregated conformation and thus biological activity. Accordingly, if an HIV Tat protein is subjected to the acid exchange reaction of Sumner-Smith *et al.*, without the presence of a reducing agent, it is expected that the HIV Tat protein would lose biological activity. Moreover, even in the presence of a reducing agent (which is not taught by Sumner-Smith *et al.*), there is no way to predict whether, and thus no expectation that, the biological activity of the HIV Tat protein would be preserved. Thus, in my judgement and opinion, if the acid exchange reaction described by Sumner-Smith *et al.* were to be applied to an HIV Tat protein purified by the RP-HPLC method of Chang *et al.*, the change in the pH of the HIV Tat protein in an aqueous environment would not be expected to preserve the biological activity of the HIV Tat protein. A subsequent lyophilization step, as disclosed in Chang *et al.*, does not entail any steps that would be expected to restore the native conformation and thus biological activity of the HIV Tat protein.

Appl. No. 09/555,534  
Attorney Docket No. 11340-003-999  
Reply to final Office Action mailed April 30, 2009

6. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that I make these statements with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of this application, and any patent issuing thereon.

Date: Oct. 26, 2009

  
\_\_\_\_\_  
Mauro Magnani, Ph.D.

## **EXHIBIT 1**

### *Curriculum Vitae – Prof. Mauro Magnani*

**MAGNANI Prof. Mauro**, Ph.D. Italian, Professor of Biochemistry

**BORN:** April 9, 1953, Italy

**LANGUAGES:** Italian, English

**EDUCATION:** Univ. Urbino, Italy, Ph.D., 1976

**PRIMARY POSITION:** Professor of Biochemistry and Director Centre of Biotechnology.

**PROFESSIONAL CAREER:** Visiting Researcher, Dept. Biochemistry, Univ. Birmingham, 1980; Visiting Prof. Dept. Biolgy, Haifa, Israel, 1983; Asst. Prof. Univ. Urbino, 1977-82, Assoc. Prof., 1982-1986, Prof. 1986 - ; Dean, Faculty of Sciences University of Urbino 1995-2001; Director Interuniversity Consortium for Biotechnology (CIB) 1998-2004; Vice Rector of the University of Urbino 2001- Sept. 24, 2009 .Include in the official list of professional biologist in Italy with n. 017484 “Ordine Nazionale Biologi”. Technical Director nominated by the “Agenzia Italiana del Farmaco, AIFA” with n. AIDT-19/2005.

**CURRENT RESEARCH:** Development of new drug delivery and drug targeting systems; Protein turnover ubiquitination and regulation of gene expression; Mechanisms of drug resistance and drug toxicity; Modulation of NF-kB and gene expression by oligonucleotide decoys, vaccine development; nanobiotechnology in drug delivery.

**PUBLICATIONS:** over 350 articles published in international refereed scientific journals; Co-editor of three books:

“*Red Blood Cell Aging*”, Plenum Press, N.Y., 1991, pp. 383.

“*The Use of Resealed Erythrocytes as Carriers and Bioreactors*”, Plenum Press, N.Y., 1992, pp. 361.

“*Erythrocyte Engineering for Drug Delivery and Targeting*”, Landes Bioscience, 2002.

**REFeree:** Programmes of the E.U.; The International Science Foundation (U.S.A.); Target Project “Biotechnology” of the National Research Council (C.N.R.); Member of the Project “Patologia clinica e terapia dell’infezione da HIV” of the Italian Ministry of Health; PRIN and FIRB Projects of Italian Ministry of University and Research; Member of Committee Post Genoma (C.N.R); Include in the “Albo degli Esperti” of M.I.U.R. and Eureka Projects of EU.

**REVIEWER:** Biotechnology and Applied Biochemistry; Nature Biotechnology; Drugs; Leukemia; European Journal Haematology; Biochimica et Biophysica Acta; Blood; Journal of Cellular Engineering; Journal of Internal Medicine; Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology; Mechanisms of Ageing and Development; Antiviral Research; Journal of Chromatography; Journal of Biological Regulators and Homeostatic Agents; Life Sciences; Biochemistry; International Journal of Biochemistry and Cell Biology; Human Gene Therapy; European Journal of Biochemistry; Clinical Pharmacokinetics; Autoimmunity; Oncogene; Haematologica; J. Controlled Release; Editorial Board: Current Drug Targets, Biotechnology.

### **PATENTS**

European Patent EP 0517986B1

M. Magnani, L. Rossi “*Transformed erythrocytes, process for preparing the same, and their use in pharmaceutical compositions*”

US Patent 5,753,221

M. Magnani, L. Rossi *“Transformed erythrocytes, process for preparing the same, and their use in pharmaceutical compositions”*

US Patent N. 6.139.836

Mauro Magnani, Ivo Panzani, Leonardo Bigi, Andrea Zanella *“Method of encapsulating biologically active agents within erythrocytes, and apparatus therefor”*.

Assignee: Dideco S.p.A., Mirandola, Italy

European Patent N. EP98830479.6

M. Magnani, G. Brandi, A. Fraternale, A. Casabianca *“Pharmaceutical composition or composition package containing a pyrimidine nucleoside analogue and a purine nucleoside analogue”*.

Brevetto C.N.R. N. RM92 A 000377

M. Magnani *“Antigeni legati alla superficie esterna di eritrociti e procedimento per la loro preparazione”*

Brevetto C.N.R. N. RM 93 A 000474

M. Magnani *“Eritrociti incorporanti alcool ossidasi e loro uso nelle intossicazioni da metanolo”*

Brevetto C.N.R.

M. Magnani, L. Rossi, G. Brandi, E. Millo, G. Damonte, U. Benatti, A. De Flora *“Profarmaco di acyclovir e suo uso in composizioni farmaceutiche”*

Brevetto di Invenzione N. MI2002A01196 – 06/06/1996 – PCT/IT 02/00368 del 13/06/2002

M. Magnani, C. Fiorucci, P. Filippone, G. Brandi, M. Paiardini. *“Derivato tetramericco dell’indol-3 carbinolo ad attività anticancerogena e metodo di sintesi del derivato stesso”*.

Brevetto di Invenzione N. TO2001A01077 – 16/11/2001

M. Magnani, F. Graziano, A. Ruzzo *“Mutazioni della linea germinale nel promotore del gene della E-caderina e metodi di diagnosi per individuare una maggiore suscettibilità al carcinoma gastrico”*.

Brevetto N. TO2003A001048 – 30/12/2003 - PCT/EP/2004/053726 – 29/12/2004

U. Benatti, G. Brandi, E. Garaci, M. Magnani, E. Millo, A.T. Palamara, L. Rossi. *“Derivati del glutathione e loro utilizzo per il trattamento di malattie virali”*.

## **EXHIBIT 2**



# HIV-1 Tat Is a Natively Unfolded Protein

## THE SOLUTION CONFORMATION AND DYNAMICS OF REDUCED HIV-1 Tat-(1–72) BY NMR SPECTROSCOPY<sup>[5]</sup>

Received for publication, October 3, 2005, and in revised form, January 11, 2006. Published, JBC Papers in Press, January 19, 2006, DOI 10.1074/jbc.M510748200

Shaheen Shojania and Joe D. O'Neil<sup>1</sup>

From the Department of Chemistry, University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada

Tat (transactivator of transcription) is a small RNA-binding protein that plays a central role in the regulation of human immunodeficiency virus type 1 replication and in approaches to treating latently infected cells. Its interactions with a wide variety of both intracellular and extracellular molecules is well documented. A molecular understanding of the multitude of Tat activities requires a determination of its structure and interactions with cellular and viral partners. To increase the dispersion of NMR signals and permit dynamics analysis by multinuclear NMR spectroscopy, we have prepared uniformly <sup>15</sup>N- and <sup>15</sup>N/<sup>13</sup>C-labeled Tat-(1–72) protein. The cysteine-rich protein is unambiguously reduced at pH 4.1, and NMR chemical shifts and coupling constants suggest that it exists in a random coil conformation. Line broadening and multiple peaks in the Cys-rich and core regions suggest that transient folding occurs in two of the five sequence domains. NMR relaxation parameters were measured and analyzed by spectral density and Lipari-Szabo approaches, both confirming the lack of structure throughout the length of the molecule. The absence of a fixed conformation and the observation of fast dynamics are consistent with the ability of Tat protein to interact with a wide variety of proteins and nucleic acid and support the concept of a natively unfolded protein.

Human immunodeficiency virus type 1 (HIV-1)<sup>2</sup>-infected resting CD4<sup>+</sup> memory T-cells form a persistent viral reservoir that is a major barrier to curing HIV-1 infection although they do not permit viral replication (1). A small, RNA-binding protein, Tat (transactivator of transcription) (2), plays a central role in the regulation of HIV-1 replication and in potential approaches to treating latently infected cells. During transcription of the viral DNA, RNA polymerase II stalls as a result of binding the negative transcription elongation factor, resulting in the production of prematurely terminated transcripts that may include the *tat* message (3). Following translation, the Tat protein is transported from the cytoplasm into the nucleus, where it binds a stem-loop structure (transactivation response element (TAR)) formed by the first 59 nucleotides of the HIV-1 RNA (4). Tat stimulates elongation of full-length transcripts by recruiting the positive transcription elongation factor b, a complex of a regulatory cyclin and cyclin-dependent kinase 9, that phosphorylates the C-terminal domain of

RNA polymerase II, components of negative transcription elongation factor, and the transcription elongation factor Spt5 (5–7). Recent results suggest that Tat activates positive transcription elongation factor b by displacing Hexim1 (hexamethylene bisacetamide-inducible protein 1) from its cyclin T1 binding site (8) and that the affinity of the Tat-cyclin T1-cyclin-dependent kinase 9 complex for TAR is regulated through Tat acetylation by histone acetyltransferases (9, 10). The absence of Tat and low levels of cyclin-dependent kinase 9 and cyclin T1 in resting CD4<sup>+</sup> T-cells are both implicated in HIV-1 latency (1). Tat may also be involved in derepression of heterochromatin, transcription initiation (11), and reverse transcription (12). In addition to its intracellular activities, the literature contains a plethora of reports of extracellular Tat activities, including general cytotoxicity, that may contribute to immune suppression (13, 14) and the development of HIV-dementia (15).

A molecular understanding of Tat activity requires a determination of its structure and interactions with cellular and viral partners (16, 17). HIV-1 Tat is a 101-residue protein encoded by two exons (3, 18). The first exon defines amino acids 1–72 that encompass an acidic and proline-rich N terminus (amino acids 1–21), a cysteine-rich region (amino acids 22–37), a core (amino acids 38–47), a basic region (amino acids 48–57), and a Gln-rich segment (amino acids 58–72) (19); it activates transcription with the same proficiency as the full-length protein (3, 20–22). Residues 1–24 form the co-activator and acetyltransferase CREB-binding protein KIX domain binding site (17). Cyclin T1 is thought to interact with the Cys-rich region of Tat (23), and mutation of any one of six of the seven Cys residues results in inactivation (3). The end of the Cys-rich region and the core are involved in mitochondrial apoptosis of bystander noninfected cells through their ability to bind tubulin and prevent its depolymerization (24). The basic region is important for TAR binding (25) and nuclear localization, and the segment between Tyr-47 and Arg-57 has been used to transport a large variety of materials, including proteins, DNA, drugs, imaging agents, liposomes, and nanoparticles across cell and nuclear membranes (26). The Gln-rich region has been implicated in mitochondrial apoptosis of T-cells (27). The second *tat* exon defines residues 73–101 and includes an RGD motif that may mediate Tat binding to cell surface integrins (28). Despite numerous attempts, the biological function of the second exon-encoded polypeptide has been difficult to determine (12, 22).

The Tat amino acid sequence has a low overall hydrophobicity and a high net positive charge, and analyses by several algorithms (29, 30) suggest that it is a natively unfolded protein (31, 32) with a possible folding nucleus between residues 42 and 75; CD spectropolarimetry confirms the lack of secondary structure (17). Proteome analysis suggests that about 33% of eukaryotic proteins are natively unfolded or contain long stretches of unfolded polypeptide, whereas unfolded proteins are much less abundant among the archaea and eubacteria (5%) (33). One advantage that unfolded proteins have over folded proteins is that they provide a much larger surface area, enabling multiple interactions with other molecules (34). The most common types of natively

\* This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada and the University of Manitoba; it was initiated with funding from the Medical Research Council of Canada and the Manitoba Health Research Council. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[5] The on-line version of this article (available at <http://www.jbc.org>) contains supplemental Figs. 15–55 and Table 15.

<sup>1</sup> To whom correspondence should be addressed. Tel.: 204-474-6697; Fax: 204-474-7608; E-mail: [joneil@cc.umanitoba.ca](mailto:joneil@cc.umanitoba.ca).

<sup>2</sup> The abbreviations used are: HIV, human immunodeficiency virus; HSQC, heteronuclear single quantum coherence; MES, 2-(N-morpholino)ethanesulfonic acid; NOE, nuclear Overhauser effect; SW, sweep width; TAR, transactivation response element; CREB, cAMP-response element.

unfolded protein in yeast are transcription factors and transcription regulators as well as kinases (35), and DNA is the most common target bound by unfolded proteins. A common mechanism of action for natively unfolded proteins involves partial or complete folding upon interaction with a binding partner (31, 32). Areas of residual structure in unfolded proteins can be critical to forming the initial site of interaction with a binding partner (32), and delineation of those sites is required for the design of drugs and antibodies that interfere with binding.

There have been several attempts to determine solution conformations of Tat and its segments both alone and in complexes.  $^1\text{H}$  NMR spectroscopy and molecular dynamics simulations suggested that Tat-(1-86) (Z-variant) forms condensed domains encompassing the core and Gln-rich regions, whereas the basic and Cys-rich regions were found to be highly flexible at pH 6.3 under reducing conditions (36). In a model of the 87-residue Tat Mal protein at pH 4.5 under oxidizing conditions, the N-terminal Trp-11 forms a hydrophobic core through interactions with Phe-38 and Tyr-47 (37). The basic region is in an extended conformation, and the Cys-rich region contains  $\beta$ -turns;  $\alpha$ -helix is found in the Gln-rich segment. A low resolution, globular conformation with some flexible segments (particularly the basic region) was deduced for  $^{13}\text{C}$ -Gly-labeled synthetic Tat-(1-86) (Bru) at pH 4.5, in the absence of reducing agents (38). An oxidized Tendamistat-Tat-(1-37) fusion protein showed multiple conformations with some evidence of helicity in the Cys-rich region (20-33) at pH 3.5 (39). A fusion protein consisting of the activation domain from the unrelated equine infectious anemia virus and Tat-(48-57) showed high helical content in the basic domain by NMR spectroscopy and CD (40).

There have also been several studies of Tat fragments in complex with TAR RNA mainly focusing on the conformation of TAR (41-45). NMR spectroscopy suggested a conformational change in Tat-(32-72) in the region of Gly-42 and Gly-44 upon binding to TAR (44).  $^1\text{H}$  NMR was also used to show that Tat-(46-55), acetylated at Lys-50, is bound in an extended conformation to the bromodomain of p300/CREB-binding protein-associated factor, a histone acetyltransferase transcriptional coactivator (10). CD spectra suggested the possibility of a conformational change in Tat-(1-86) upon binding to the KIX domain of CREB-binding protein (17).  $^{15}\text{N}$  NMR relaxation measurements showed that Tat-(47-58) becomes slightly more ordered on binding heparin (46), and CD studies of overlapping peptide fragments suggested that the most flexible regions are those that are adjacent to the basic region (47).

Determining the mechanisms by which proteins attain their biologically active three-dimensional structures is a fundamental goal of biology, with implications ranging from molecular recognition to understanding amyloid diseases (48). Multinuclear NMR spectroscopy is yielding insights into the folding process by characterizing unfolded proteins and equilibrium intermediates. The detection of residual native and nonnative structures in strongly denatured proteins suggests that they are important in seeding the early steps of protein folding (49-56). Molten globule states of proteins have also been investigated by NMR and provide insight into the nature of the folding pathway and the structure of the transition state (57, 58). Thus, although NMR spectroscopy is not ideally suited to kinetic studies of protein folding, it is incomparable in its ability to provide detailed information about the structures and dynamics of unfolded and partially folded proteins (59).

$^1\text{H}$  NMR spectra of Tat are very difficult to assign due to poor dispersion of the resonances (36-39). In addition, fully and partially oxidized material is likely to be structurally heterogeneous, including the formation of disulfide cross-linked oligomers often observed on polyacrylamide gels and adding to the difficulty in structural characterization. To increase the dispersion of the NMR signals and permit dynamics analysis by multinuclear NMR spectroscopy, we have prepared uniformly

$^{15}\text{N}$ - and  $^{15}\text{N}/^{13}\text{C}$ -labeled Tat-(1-72) protein. The protein is unambiguously reduced at pH 4.1, and NMR analysis provides insight into the intracellular folding and dynamics of this intrinsically disordered protein and the folding propensities of its five sequence domains.

## EXPERIMENTAL PROCEDURES

**Preparation of Histidine-tagged Tat**—The detailed Tat-(1-72) expression and purification protocol will be published elsewhere. Briefly, an *Escherichia coli* codon-optimized synthetic exon 1 *tat* gene contained in pSV2tat72 was obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, National Institutes of Health, from Dr. Alan Frankel (60) and cloned into pET28b(+) (Novagen, Madison, WI), enabling protein expression with an N-terminal His<sub>6</sub> segment and T7 epitope that adds 20 residues to the 72-residue protein. Protein expression was from *E. coli* BL21(DE3)plysS cells (Novagen, Madison, WI) in M9 minimal medium supplemented with either  $^{15}\text{NH}_4\text{Cl}$  or  $^{15}\text{NH}_4\text{Cl}$  and U- $^{13}\text{C}_6$ -D-glucose (Cambridge Isotope Laboratories Inc., Andover, MA). The expression protocol was modified from recently published methods (61) that reduce the consumption of isotopically labeled ingredients. The protein was purified by elution from a Talon<sup>TM</sup> (Cobalt-Superflow) metal affinity resin (Clontech, Palo Alto, CA) in a gravity flow column in 50 mM sodium acetate buffer, 6 M guanidine HCl, and 10 mM Tris(2-carboxyethyl)phosphine (Sigma) adjusted to pH 4.0 with sodium hydroxide. The eluted protein was pooled and serially dialyzed against degassed 0.1, 0.05, and 0.01 M ammonium acetate at pH 3 for ~6 h each. A final dialysis was done against degassed water for 4 h. The dialysate was then frozen and freeze-dried.

**NMR Sample Preparation**—Freeze-dried protein was dissolved in degassed buffer containing 50 mM acetate- $d_4$ , 20 mM MES buffer, 80  $\mu\text{M}$  sodium sulfite, 0.02% sodium azide, and 5%  $\text{D}_2\text{O}$ , pH 4.0. The resulting protein solution had a measured pH of 4.1, uncorrected for isotope effects. The sample was placed in an NMR tube that had been purged with argon for 15 min, and the dissolved protein was added to the sample tube under an argon atmosphere. The NMR tube cap was sealed with Teflon tape (Dupont Dow Elastomers LLC, Wilmington, DE). The final protein concentration was about 1 mM. The purity of the protein was confirmed by matrix-assisted laser desorption ionization mass spectrometry. The pH titration was done by dissolving separate protein samples in degassed buffers as described above, with subsequent verification of the pH.

**NMR Backbone Assignments**—All backbone assignment experiments for the 92-residue  $^{13}\text{C}/^{15}\text{N}$ -labeled histidine-tagged Tat-(1-72) were done on a 600-MHz Varian INOVA spectrometer equipped with a triple resonance probe head at 20.2  $^\circ\text{C}$ , using standard Varian BioPack pulse sequences (62-67) (see Table 1). The NMR probe was calibrated with methanol (68), and all spectra were processed with NMRPipe (69). Spectra were apodized using a squared cosine bell function, zero-filled to twice the data set size, and linear predicted (forward-backward with eight prediction coefficients) prior to Fourier transformation. The dimensions of the resulting processed data sets were  $4096 \times 1024$  for the  $^1\text{H}/^{15}\text{N}$  HSQC experiment and  $2048 \times 256 \times 128$  for all three-dimensional experiments. The pulse sequences used are sensitivity-enhanced (with the exception of the HNHA experiment) and use gradients for coherence selection and water suppression (70). Radiation damping was suppressed with a water flip-back pulse (1.42 ms).  $^{15}\text{N}$  decoupling during acquisition was done using the WALTZ-16 sequence (71) with a 7.196-kHz field strength.  $^1\text{H}$  chemical shifts were referenced to the water signal that resonates 4.821 ppm from 2,2-dimethyl-2-silapentane sulfonate at 293 K (68).  $^{15}\text{N}$  and  $^{13}\text{C}$  referencing were done indirectly relative to 2,2-dimethyl-2-silapentane sulfonate as recommended (72).

TABLE 1

## Acquisition parameters for the NMR experiments

All experiments acquired with a 0.7-s postacquisition relaxation delay.

| Experiment                                | Transients | Complex points | SW[ <sup>1</sup> H] | SW[ <sup>13</sup> C] | SW[ <sup>15</sup> N] |
|---|------------|----------------|---------------------|----------------------|----------------------|
|   |            |                |                     | ppm                  |                      |
| <sup>1</sup> H/ <sup>15</sup> N HSQC (62) | 32         | 2048 × 256     | 12                  |                      | 36                   |
| HNCACB (63)                               | 16         | 1024 × 64 × 32 | 12                  | 70                   | 30                   |
| CBACA(CO)NH (64)                          | 8          | 1024 × 64 × 32 | 10                  | 70                   | 24                   |
| HNCO (65)                                 | 8          | 1024 × 64 × 32 | 10                  | 8                    | 24                   |
| HN(CA)CO (66)                             | 8          | 1024 × 64 × 32 | 10                  | 8                    | 24                   |
| HNHA (67)                                 | 8          | 1024 × 64 × 32 | 10                  | 10                   | 30                   |

Random coil values were subtracted from the measured chemical shifts after correction for sequence effects as described in Ref. 73. The  $\delta_{\text{HNH}\alpha}^{\text{measured}}$  were not corrected for relaxation effects and are likely to be 5–10% underestimated (68). The  $\delta_{\text{HNH}\alpha}^{\text{calculated}}$  values were calculated by subtracting the measured COIL values reported in (74).

**NMR Relaxation Measurements**—NMR relaxation data were collected on both <sup>15</sup>N-labeled and <sup>13</sup>C/<sup>15</sup>N-labeled His-tagged Tat-(1-72) on a Varian INOVA 600 MHz spectrometer at the University of Manitoba and on a Varian INOVA 800-MHz spectrometer at the University of Alberta (NANUC) with triple resonance probe heads at 20 °C, using BioPack pulse sequences (70, 75, 76). Cross-peak intensities were measured as peak heights. Spectra were processed with NMRPipe (69), which was also used to fit the relaxation data to two-parameter decays. The errors in the relaxation rates were calculated using the signal/noise ratios of the individual peaks and the fits of the data to the decays. Duplicate measurements were made to verify the error estimates.

A total of nine data sets were acquired to obtain  $R_1$  relaxation rates using relaxation delays of 0, 50, 100, 250, 500, 1000, 1500, 3000, and 4000 ms.  $R_{1\rho}$  measurements were made with eight data sets with spin lock times of 30, 60, 90, 120, 150, 180, 210, and 240 ms. The <sup>15</sup>N spin lock continuous wave field strength for the  $R_{1\rho}$  relaxation experiments was 1.5 kHz, with 90° pulse lengths of 166.755 and 125.029 ms for the 600 and 800 MHz fields respectively. The  $R_{1\rho}$  measurements were corrected for offset from the carrier, using the measured  $R_1$  values as described in Ref. 76. The peaks at the outer edges of the spectra required correction by less than 10%.  $R_2$  measurements were done at 600 MHz only and with Carr-Purcell-Meiboom-Gill times of 30, 60, 90, 120, 150, 180, 210, and 240 ms. The  $R_1$  and  $R_{1\rho}$  data were acquired using four transients, whereas eight transients were collected for the  $R_2$  experiments; the post-acquisition relaxation delay was 5 s. Data collected at 600 MHz were 2048 × 256 complex points with SW[<sup>1</sup>H] = 10 ppm and SW[<sup>15</sup>N] = 24 ppm. The steady-state <sup>1</sup>H-<sup>15</sup>N NOE values were obtained from ratios of peak heights from experiments with ( $I_{\text{NOE}}$ ) and without ( $I_{\text{noNOE}}$ ) saturation of the protons for 5 s at the beginning of the experiment. The heteronuclear NOE values were then obtained from  $(I_{\text{NOE}} - I_{\text{noNOE}})/I_{\text{noNOE}}$ . The spectra were acquired with 32 transients, a 5-s relaxation delay, and the same resolution as in the  $R_1$  and  $R_{1\rho}$  experiments. Water suppression was achieved through the use of gradients to select for the <sup>15</sup>N-<sup>1</sup>H coherence (70). Data collection at 800 MHz was done exactly as at 600 MHz, but the resolution for the experiments was 672 × 256 complex points with SW[<sup>1</sup>H] = 15 ppm and SW[<sup>15</sup>N] = 26 ppm.

**Relaxation Data Analysis**—The measurement of NMR relaxation rates provides a window on protein dynamics over a broad range of time scales: <sup>15</sup>N longitudinal ( $R_1$ ), transverse ( $R_2$ ), rotating frame ( $R_{1\rho}$ ), and heteronuclear cross-relaxation (contained in the NOE) rates are sensitive to dynamics on the picosecond to nanosecond time scales, and  $R_2$  and  $R_{1\rho}$  can also be sensitive to conformational exchange ( $R_{\text{ex}}$ ) on the millisecond to microsecond time scales. The equations relating the macroscopic rates of relaxation ( $R_i$ ) to the values of the spectral density

of motions ( $J$ ) at the nuclear spin transition frequencies ( $\omega$ ) were given by Abragam (78) but can be approximated as follows (79, 80).

$$R_1 = D(3J(\omega_N) + 7J(0.921\omega_H)) + C/J(\omega_N) \quad (\text{Eq. 1})$$

$$R_2 = R_{1\rho} = D(4J(0) + 3J(\omega_N) + 13J(0.955\omega_H)) + (C/6)(4J(0) + 3J(\omega_N)) + R_{\text{ex}} \quad (\text{Eq. 2})$$

$$\text{NOE} = (\gamma_H/\gamma_N)D(5J(0.87\omega_H)/R_1) \quad (\text{Eq. 3})$$

The constants in Equations 1–3 are defined by  $D = (\mu_0/4\pi)^2 \times (\gamma_H^2\gamma_N^2\hbar^2)/4r_{\text{NH}}^6$  and  $C = (\Delta^2\omega_N^2)/3$ , where  $\mu_0$  is the permeability constant of free space ( $4\pi \times 10^{-7} \text{ kg}\cdot\text{m}\cdot\text{s}^{-2}\cdot\text{A}^{-2}$ ),  $\gamma_H$  is the proton magnetogyric ratio ( $2.68 \times 10^8 \text{ rad}\cdot\text{s}^{-1}\cdot\text{T}^{-1}$ ),  $\gamma_N$  is the magnetogyric ratio of <sup>15</sup>N ( $-2.71 \times 10^7 \text{ rad}\cdot\text{s}^{-1}\cdot\text{T}^{-1}$ ),  $r_{\text{NH}}$  is the proton-nitrogen internuclear separation (104 pm),  $\Delta$  is the difference between the parallel and perpendicular components of the <sup>15</sup>N chemical shift tensor ( $-163 \text{ ppm}$ ), and  $\hbar$  is Planck's constant divided by  $2\pi$  ( $1.05 \times 10^{-34} \text{ J}\cdot\text{s}$ ). Note that  $R_2$  and  $R_{1\rho}$  are determined by the same combination of spectral density values as long as the <sup>15</sup>N spin lock is on resonance for all spins (81). An exact solution of Equations 1–3, which contain six unknowns, is not possible, but because the spectral density is relatively flat at high frequencies, the number of variables can be reduced to three by estimating the spectral density values near  $J(\omega_H \pm \omega_N)$ , as described by Farrow *et al.* (82). Briefly,  $J(0.87\omega_H)$  is determined from Equation 3, and  $J(0.921\omega_H)$  and  $J(0.955\omega_H)$  are calculated directly from it using the assumption that  $J(\omega) \propto 1/\omega^2$ . One advantage to measuring  $R_{1\rho}$  is that, in contrast to  $R_2$ , contributions from conformational exchange are eliminated ( $R_{\text{ex}} \sim 0$ ) as long as the nitrogen carrier is placed on resonance and the spin lock power is sufficiently high (51, 83, 84). This allows a direct calculation of  $J(\omega_N)$  and  $J(0)$  (strictly  $J(\omega_0)$ , the magnitude of the effective field/frequency in the presence of the spin lock) from the measured relaxation rates and steady-state NOE. Uncertainties in the spectral densities were determined by repeating the calculations 500 times using the S.D. values of the NMR measurements and Monte Carlo methods to generate a normal distribution and are described in more detail in Ref. 85. The calculations were done using the program *Mathematica* 5.0 and its simulated annealing protocol (86); statistical analyses were done with the program JMP IN 5.1.

Relaxation measurements were done at two fields to permit a finer mapping of the spectral density and more specifically to test the assumptions inherent in the reduced spectral density analysis. In addition, since  $R_{\text{ex}}$  scales with the square of the applied magnetic field, it is possible to determine the contribution of  $R_{\text{ex}}$  to Equation 2 by measuring relaxation parameters at two fields. Thus,  $R_{\text{ex}}$  and  $J(0)$  values were calculated from the relaxation measurements at 600 and 800 MHz as described in Ref. 82. The  $R_{1\rho}$  relaxation data were modeled by assuming

that the effect of its neighbors ( $j$ ) on the correlation time of a residue ( $i$ ) decreases exponentially as the distance from the residue increases and was first described in Ref. 51,

$$R_{1\rho}(i) = R_{1\rho}^{\text{int}} \sum_{j=1}^N V \exp\left(-\frac{|(i-j)|}{\lambda}\right) \quad (\text{Eq. 4})$$

where  $R_{1\rho}^{\text{int}}$  is an intrinsic residue relaxation rate,  $N$  is the length of the polypeptide,  $V$  is the residue molecular volume (87), and  $\lambda$  is the persistence length of the polypeptide in residues.

A different solution to Equations 1–3 was proposed by Lipari and Szabo (88), who derived a simplified spectral density function  $J(\omega)_{\text{LS}}$  on the assumption that global molecular reorientation ( $\tau_c$ ) and fast internal motions ( $\tau_e$ ) are stochastically uncorrelated (89),

$$J(\omega)_{\text{LS}} = \left(\frac{2}{5}\right) \left( \frac{S^2 \tau_c}{1 + (\omega \tau_c)^2} + \frac{(1 - S^2) \tau_e}{1 + (\omega \tau_e)^2} \right) \quad (\text{Eq. 5})$$

where  $(1/\tau) = (1/\tau_c) + (1/\tau_e)$ . The Lipari-Szabo “model-free” spectral density reduces the number of unknown parameters in Equations 1–3 to three:  $S^2$ , the square of the generalized order parameter, which indicates the degree of spatial freedom of the internal motion;  $\tau_c$ , the global rotational correlation time for molecular reorientation, and  $\tau_e$ , the effective internal rotational correlation time, which is related to both the amplitude and the rate of internal motion. The separability of internal and overall dynamics is questionable for a random coil polymer, but comparisons of the Lipari-Szabo parameters with those obtained for other folded and unfolded proteins can be instructive. We have implemented the data analysis approach developed by Schurr *et al.* (90), in which all three Lipari-Szabo parameters are optimized for each residue individually, since this is reported to provide a significantly better fit to the NMR data (90). The errors in the Lipari-Szabo parameters were determined by Monte Carlo analysis as described above for the spectral density analysis, except that only 100 points were calculated (85).

## RESULTS AND DISCUSSION

**NMR Resonance Assignments**—Sequential assignments of  $^1\text{H}^N$ ,  $^{15}\text{N}$ ,  $^{13}\text{C}'$ ,  $^{13}\text{C}^\alpha$ ,  $^1\text{H}^\alpha$ , and  $^{13}\text{C}^\beta$  resonances were done entirely with three-dimensional heteronuclear triple resonance experiments that use one- and two-bond scalar couplings to connect the atoms (68). The experiments take advantage of the comparatively wide chemical shift dispersions of  $^{15}\text{N}$ ,  $^{13}\text{C}'$ , and  $^1\text{H}^N$  resonances in unfolded proteins (91, 92). All of the backbone resonances were sequentially assigned, except for Met-1, Met-21, Phe-52, Phe-58, Arg-77, and Pro-78. The Met-1 amino and Gly-2 amide protons exchange too rapidly to be observed (see below). Resonances from Phe-52/58 could be assigned to residue type only, because they are both preceded by weak Cys resonances. Arg-77 and Pro-78 are part of the difficult sequence Arg-Arg-Arg-Pro-Pro and could not be unambiguously sequentially assigned. Of the assigned Pro resonances, all but one have  $^{13}\text{C}^\beta$  chemical shifts characteristic of the *trans* peptide. Pro-38  $^{13}\text{C}^\beta$  shifts are outside the canonical chemical shifts for the *trans* configuration but are nearer the *trans* than the *cis* configuration. As an example, parts of the HN(CA)CO and HNCACB experiments used for the backbone assignment of residues Gln<sup>83</sup>–Leu<sup>89</sup> are shown in Fig. 1. The assignments are listed in Table 1 of the supplemental data.

A  $^1\text{H}/^{15}\text{N}$  HSQC spectrum of Tat is shown in Fig. 2 and contains three notable features. First, the cross-peaks are separated into regions typical of the residue random coil positions (Gly, Ser/Thr, other backbone amides), and the chemical shift dispersions of  $^{15}\text{N}$  (20 ppm) and  $^1\text{H}^N$  (1.1 ppm)

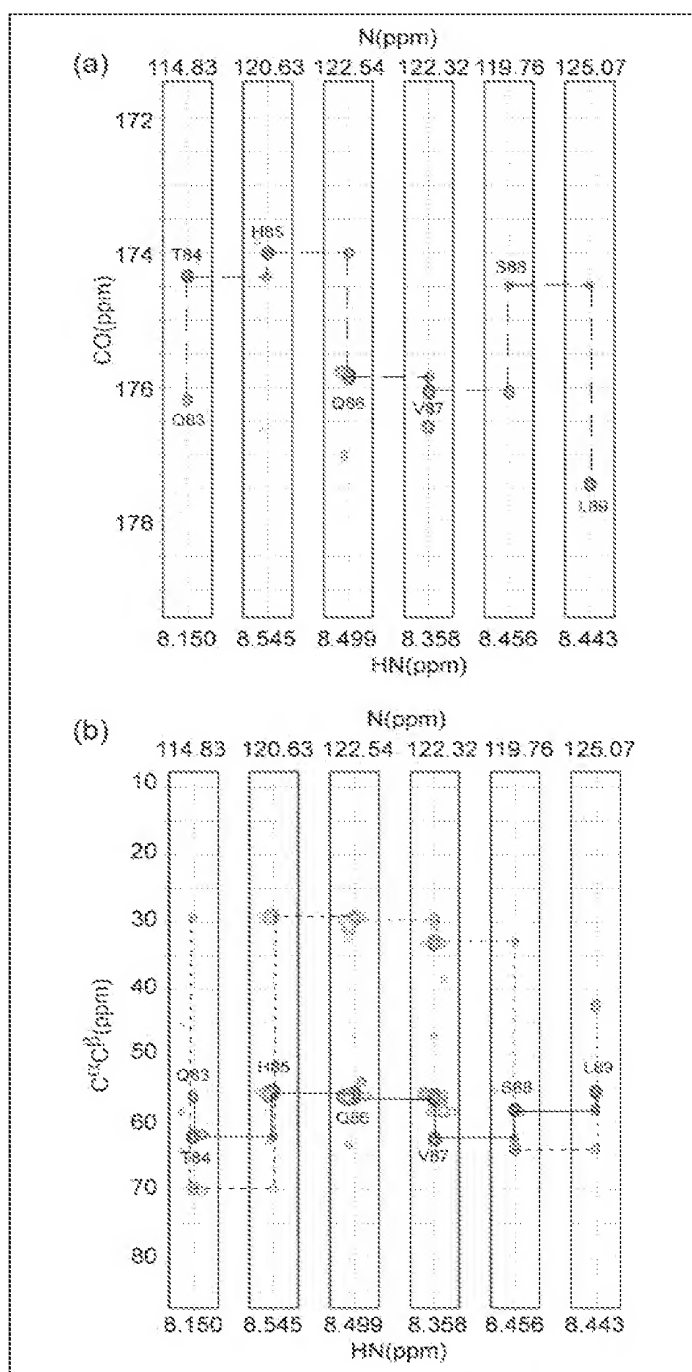


FIGURE 1. Strip plots extracted from three-dimensional, amide-detected heteronuclear NMR experiments for backbone assignment. Inter- and intraregional correlations are obtained from (a) an HN(CA)CO spectrum correlating  $\text{H}^N(i)$  and  $\text{N}(i)$  with  $\text{C}'(i)$  and  $\text{C}'(i-1)$  resonances and (b) an HNCACB spectrum correlating  $\text{H}^N(i)$  and  $\text{N}(i)$  with  $\text{C}^\alpha(i)$ ,  $\text{C}^\alpha(i-1)$ ,  $\text{C}^\beta(i)$ , and  $\text{C}^\beta(i-1)$  resonances. A segment of the histidine-tagged Tat-(1-72) at pH 4.1 and 293 K is shown depicting connectivity between residues 83 and 89. Correlations in a are shown with long dashed lines; in b, the  $\text{C}^\alpha$  correlations are connected with solid lines, and correlations of  $\text{C}^\beta$  are connected with short dashed lines.

resonances are virtually identical to those of strongly denatured ubiquitin (93). Second, the peaks exhibit a range of intensities, with nearly all of the weak and medium intensity cross-peaks falling in the sequence between Cys-47 and Leu-63. This suggests that this region of the protein is under-

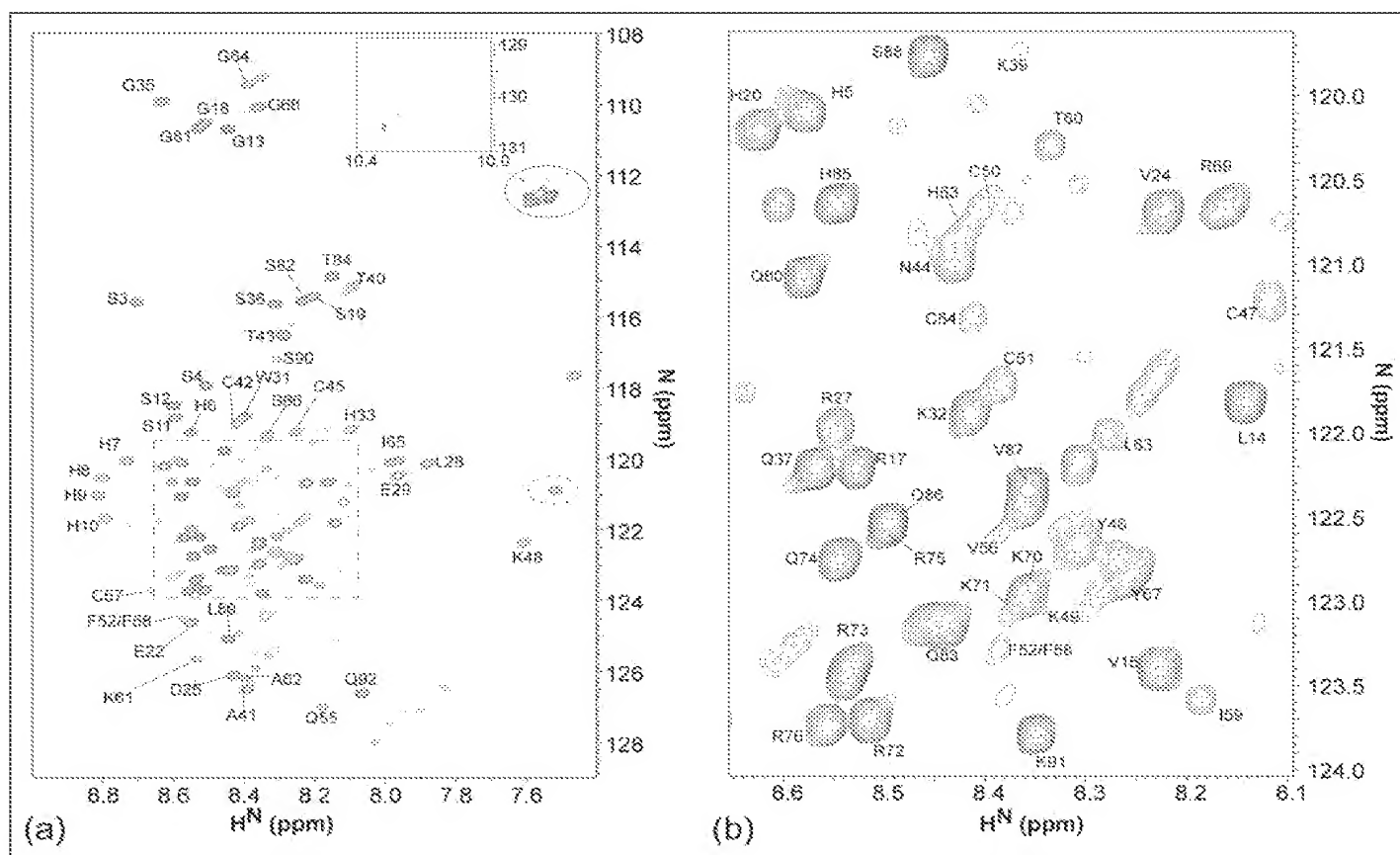


FIGURE 2. Two-dimensional  $^1\text{H}/^{15}\text{N}$  HSQC spectrum of Tat at pH 4.1 and 293 K. *a*, backbone amide proton resonances are narrowly dispersed over the range of 7.4–8.9 ppm. The inset of *a* shows the three peaks associated with the side chain of the single Trp residue. *b*, the dashed region of *a* is expanded to allow labeling of the crowded central region of the spectrum. Backbone resonance assignments of 80 of the 84 nonproline resonances are shown. Side chain Asn and Gln  $\text{NH}_2$  resonances are outlined with a solid ellipse, and side chain Arg resonances are outlined in a dashed ellipse.

going conformational exchange on the millisecond to microsecond time scale, indicating transient structure formation.

Third, the spectrum contains several more mostly weak cross-peaks than can be accounted for by the backbone and side chain atoms of a 92-residue protein. Sixteen of the additional resonances have been sequentially assigned (indicated with the designation "a" in supplemental Table 1S), some were assignable to amino acid type, and some could not be unambiguously assigned. Eleven of the assigned minor resonances fall in the sequence between Cys-45 and Arg-69. One example of the multiplicity of cross-peaks is shown for the single Trp at position 31 (see Fig. 2*a*, inset), which exhibits one strong and two weaker side chain indole amine cross-peaks. The Trp is preceded by a Pro, so one of the two minor peaks could arise from the *cis*-peptide bond isomer. Another possible explanation is that some of the minor peaks are due to the presence of minor amounts of oxidized Cys residues that are either unassigned or unobserved. Another interesting example is Gly-64, which exhibits two amide cross-peaks of approximately equal intensity (Fig. 2). The nearest Pro to Gly-64 is separated from it by 14 residues, ruling out *cis-trans* Pro isomerism as an explanation of the peak multiplicity. This suggests that some segments of the reduced, monomeric Tat protein exist in multiple conformations that are in slow equilibrium on the chemical shift time scale (milliseconds to seconds). In the case of Gly-64, the two resonances have comparable intensity, suggesting equal populations of 2 conformers, whereas in many other cases one

resonance is significantly more intense than those arising from alternate conformers, suggesting one dominant conformer and minor alternates. Gly-64 is located between Leu and the  $\beta$ -branched Ile, and steric crowding could locally restrict the dynamics of the Gly amide. Since there is a variation in the intensities of the duplicate peaks across the sequence, this immediately suggests that the conformers populated arise from local interactions, as expected in an unfolded protein.

**Chemical Shifts and  $^3J_{\text{HNH}\alpha}$ .**—The NMR chemical shift is a sensitive indicator of conformation, and assignment of backbone chemical shifts permits an analysis of secondary structure by comparison with random coil values corrected for local sequence effects (73, 94). Consensus multinuclear ( $^{13}\text{C}'$ ,  $^{13}\text{C}^\alpha$ ,  $^{13}\text{C}^\beta$ , and  $^1\text{H}^\alpha$ ) chemical shift indexing (95) suggests that all segments of the reduced protein at pH 4.1 exist in a random coil conformation (data not shown). Examination of the individual chemical shift indexing plots shown in Fig. 3 indicates that a majority of the resonances are within the random coil range and that rarely are there more than three consecutive resonances in the  $\alpha$ -helix or  $\beta$ -sheet chemical shift ranges. However, among the  $^1\text{H}^\alpha$ ,  $^{13}\text{C}^\alpha$ , and  $^1\text{H}^\alpha$  resonances, there appears to be a slight weighting of the conformations by the  $\alpha$ -helix, the most consistent classification being for the segment around Glu-29. Unlike some other denatured proteins (50, 51, 96), there is less evidence of a tendency to the  $\beta$ -sheet conformation, perhaps because of a paucity of hydrophobic  $\beta$ -branched amino acids in Tat-(1-72) (two Ile and four Val). However, there is a slight suggestion of this in the region between Val-56 and Ile-59. The conclusions based



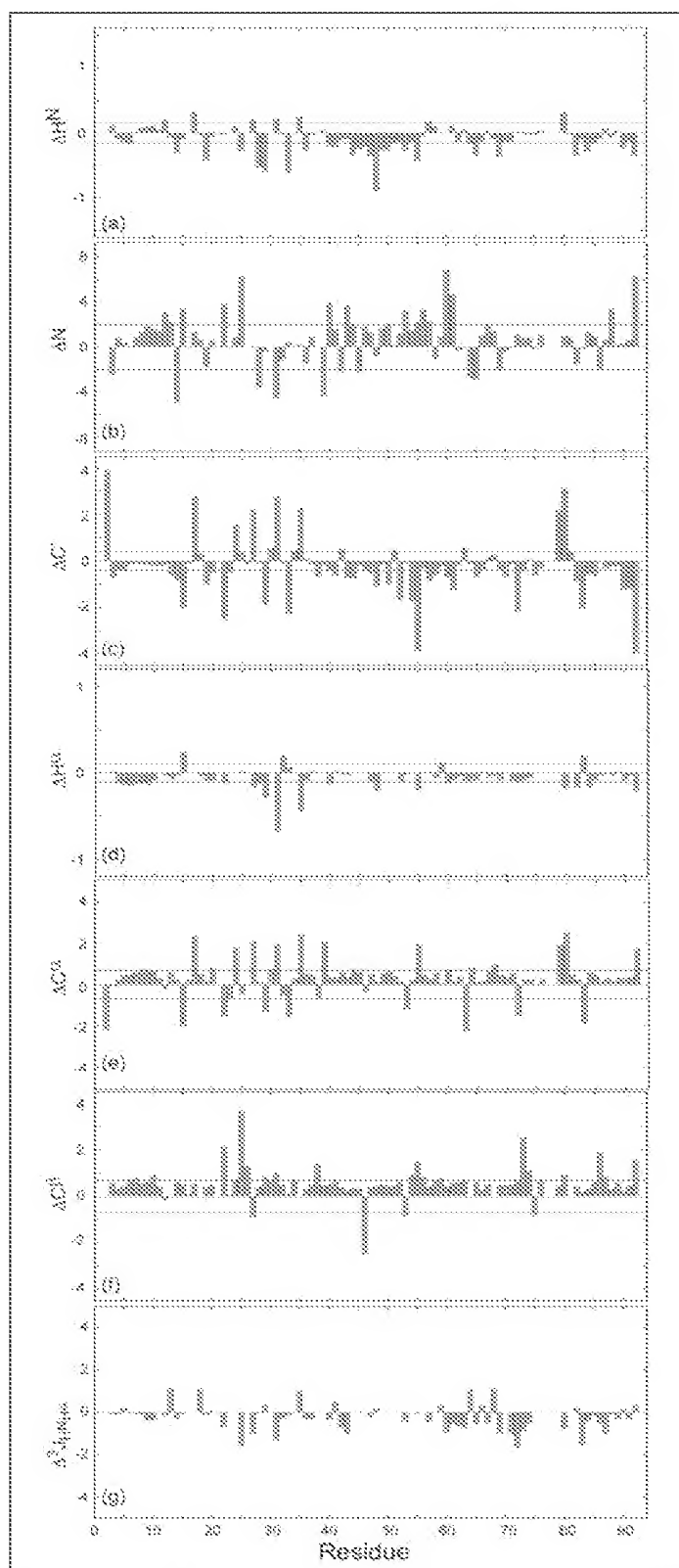


FIGURE 3. Chemical shift difference index plots of  $H^N$  (a),  $N$  (b),  $C'$  (c),  $H^\alpha$  (d),  $C^\alpha$  (e), and  $C^\beta$  (f), g, difference plot of the  $^2J_{HNH\alpha}$  coupling constant from random coil coupling constant. The random coil values for  $H^N$ ,  $N$ ,  $C'$ ,  $C^\alpha$ , and  $H^\alpha$  have been adjusted for sequence dependence (correction factors for  $C^\beta$  are unavailable). Reference lines in plots a–f correspond to thresholds where differences begin to reflect regions of possible

on chemical shift indexing are supported by the uncorrected  $^3J_{HNH\alpha}$  measurements, which are all in the range of 5.5–7.1 Hz with a mean value of 6.71 characteristic of unfolded molecules (74). The results are shown in Fig. 3g, in which the differences of the measured values from random coil values are presented. They show that the entire polypeptide is undergoing rapid sampling of the  $\alpha$ -helix and  $\beta$ -sheet regions of Ramachandran space with a slight preference for the  $\alpha$ -helix in most segments.

**NMR Relaxation**—Despite the relatively poor chemical shift dispersion in the NMR spectra, relaxation data could be measured for 64/84 (60/84) non-Pro resonances at 600 MHz (800 MHz). As indicated in Fig. 4a, the steady-state heteronuclear NOEs measured at 600 and 800 MHz exhibit a relatively featureless, flattened bell-shaped variation with amino acid sequence, as expected for an unfolded protein (97). The NOEs range from  $-3.3$  ( $-2.6$ ) to  $-0.60$  ( $-0.41$ ) with mean values of  $-1.27$  ( $-0.933$ ) at 600 MHz (and 800 MHz). For comparison, an average NOE of about  $-0.2$  is observed for several folded proteins with similar lengths of polypeptide chain (98–100). The more negative NOE values for Tat indicate much less restricted dynamics on the nanosecond to picosecond time scales than for the folded proteins. The ends of Tat exhibit the most negative values indicative of faster dynamics and the values gradually increase away from the C terminus, whereas the increase away from the N terminus is steeper. Significant deviations from the average values are observed for Thr-43 and Lys-61-Ala-62.

The  $R_1$  measurements show a similar bell-shaped profile with two notable features (Fig. 4b). The  $R_1$  values range from  $0.75\text{ s}^{-1}$  ( $0.89\text{ s}^{-1}$ ) to  $1.91\text{ s}^{-1}$  ( $1.92\text{ s}^{-1}$ ) with mean values of  $1.46\text{ s}^{-1}$  ( $1.43\text{ s}^{-1}$ ) at 600 MHz (and 800 MHz). Several folded proteins of similar size show slightly higher average  $R_1$  values on the order of  $1.5\text{ s}^{-1}$  at 600 and 750 MHz (99, 101). The slower relaxation in Tat indicates a shorter rotational correlation time and faster dynamics on the nanosecond to picosecond time scale than for a folded protein. Similar to the NOE values, the  $R_1$  rates decline near the ends of the protein more steeply at the N terminus than the C terminus. The lowest rates, apart from the termini, are found in the segment connecting the Cys-rich and basic regions, between Thr-60 and Ser-66, and suggest fast dynamics there.

The rotating frame longitudinal relaxation rates ( $R_{1\rho}$ ) measured for Tat at 600 and 800 MHz are plotted in Fig. 4c. The  $R_{1\rho}$  rates range from  $1.5\text{ s}^{-1}$  ( $1.3\text{ s}^{-1}$ ) to  $5.9\text{ s}^{-1}$  ( $7.2\text{ s}^{-1}$ ) with mean values of  $3.26\text{ s}^{-1}$  ( $3.29\text{ s}^{-1}$ ) at 600 and 800 MHz, respectively. The  $R_2$  rates, measured at 600 MHz (supplemental data), range from  $1.6$  to  $7.1\text{ s}^{-1}$ , with an average value of  $3.5\text{ s}^{-1}$ . The differences between the  $R_2$  and the  $R_{1\rho}$  rates presumably arise from contributions to the former from slow conformational exchange, but hydrogen exchange could also contribute to the higher  $R_2$  values. In general, low  $R_{1\rho}$  and  $R_2$  measurements indicate unrestricted fast dynamics, whereas high values suggest restricted fast dynamics and possible contributions from slow conformational exchange (100). In folded proteins of similar length to Tat, the  $R_2$  values in the absence of exchange are on the order of  $8\text{ s}^{-1}$  (99). The low  $R_{1\rho}$  and  $R_2$  values and the negative NOE values measured for Tat indicate large amplitude fluctuations on the picosecond-nanosecond time scale characteristic of a random coil conformation.

The  $R_{1\rho}$  relaxation data at 600 MHz were fit to Equation 4, in which the influence of neighboring residues is modeled as a decaying exponential (51). The flattened, bell-shaped solid curve shows the behavior predicted for a random-coil polymer of uniform composition (Fig. 4d). The

secondary structure formation. The plot ranges in a–f correspond to two S.D. values from the mean value determined from the chemical shift tables in the Biological Magnetic Resonance Bank data base.

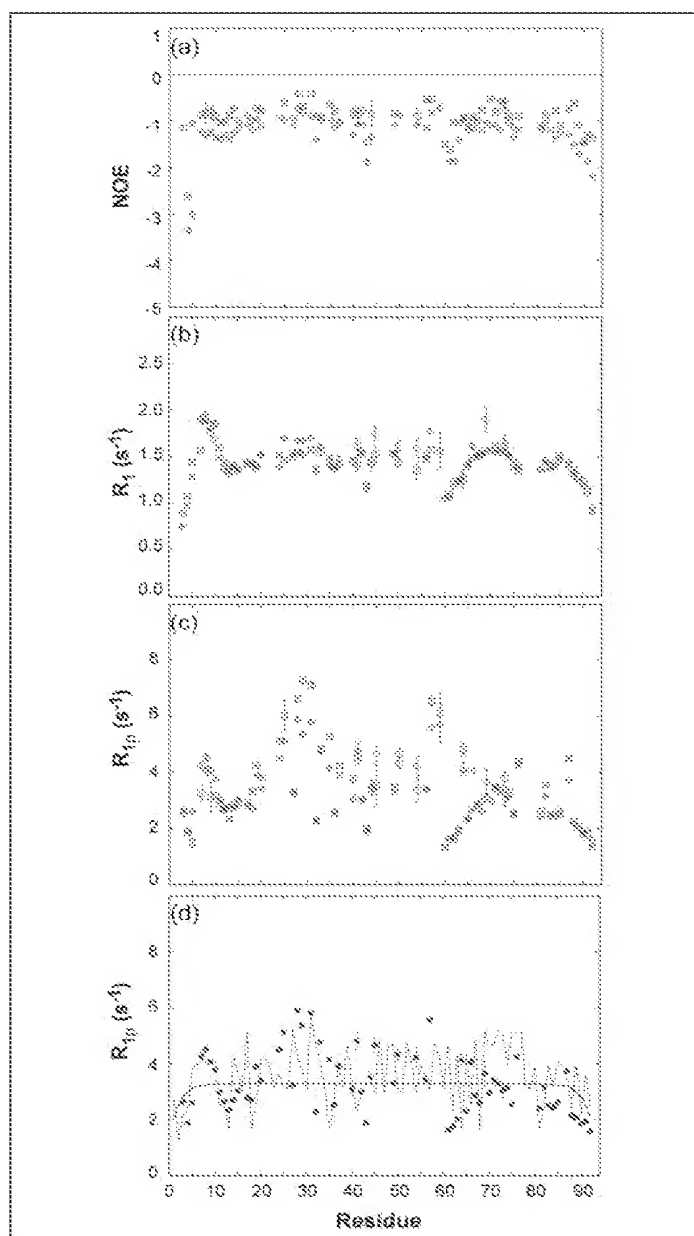


FIGURE 4. Relaxation measurements of Tat-(1-72) protein at pH 4.1 and 293 K, determined at 600 MHz ( $\circ$ ) and 800 MHz ( $\square$ ) field strengths. *a*, heteronuclear NOE; *b*, longitudinal relaxation,  $R_1$ ; *c*, rotating frame relaxation,  $R_{1\rho}$ ; *d*,  $R_{1\rho}$  data at 600 MHz field strength ( $\bullet$ ) plotted along with the predicted behavior for a random coil polymer of uniform composition (solid line) and the variation in relaxation when residue contributions are weighted by residue volume (dashed line).

dashed line shows the variation in relaxation when residue contributions are weighted by residue volume (87). Although a number of individual residues deviate from the volume-weighted model, overall the theoretical line follows the data fairly closely and is a much better fit than the uniform polymer model. The minima in the model correspond mainly to the small, flexible residues Gly, Ala, and Ser, whereas the maxima are found at the positions of Trp, Arg, and Lys. In contrast to some other applications of this model to denatured proteins (51, 102, 103), there are no obvious regions with large positive deviations from the theoretical curve, further evidence that reduced Tat-(1-72) at pH 4.1 is predominantly a random coil. The segment from Pro-23-Pro-38

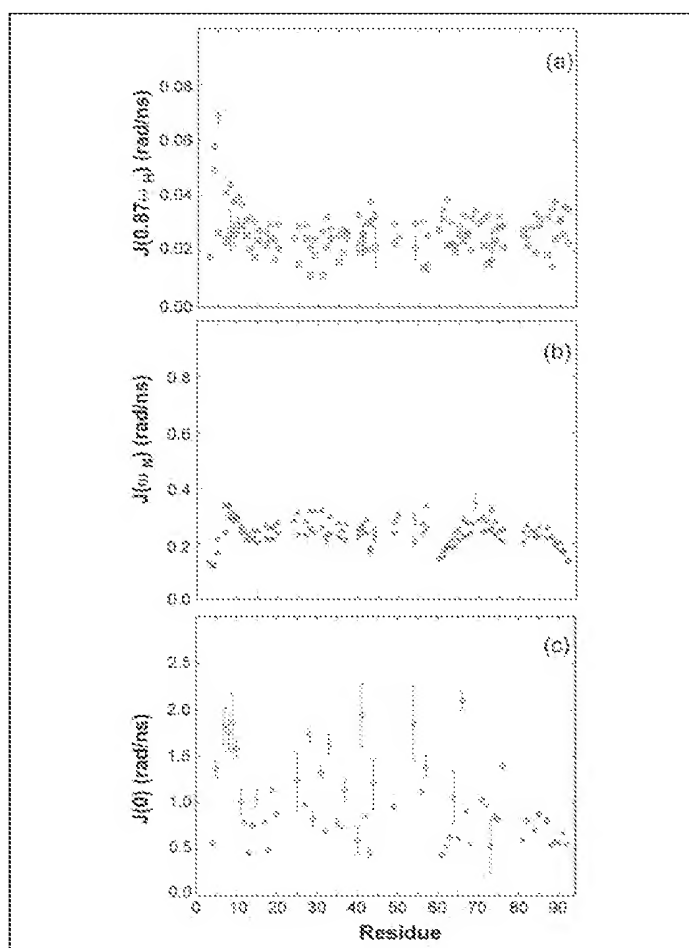


FIGURE 5. Reduced spectral density mapping of motions for His-tagged Tat-(1-72) at pH 4.1 and 293 K. *a*, at 0 MHz ( $\diamond$ ), calculated using the measurements from two fields according to the method of Ref. 82; *b*, at 61 MHz ( $\triangle$ ) and 81 MHz ( $\nabla$ ); *c*, 522 MHz ( $\circ$ ) and 696 MHz ( $\square$ ), based on estimation of  $J(0.87\omega_H)$ .

contains 5 prolines, and the measured  $R_{1\rho}$  values for most residues in this region are greater than the calculated ones. This suggests that the prolines restrict dynamics on the millisecond to microsecond time scale, stiffening the backbone in this region. In the C terminus, from residue 60 onward, the measured values generally fall below the calculated ones, suggesting greater flexibility in this region of the molecule. One exception is the high value for Gly-64, suggesting restricted motion and slow exchange at this position (102).

One final observation is that the region of the protein spanning residues 45–60 (Cys-rich region and core) contains the fewest dynamics measurements. This is because the peak intensities in this region are low. Since the largest number of assigned minor peaks is also found in this segment (supplemental data), this supports the suggestion that some residues in this segment undergo slow conformational exchange and are the most likely sites of folding nuclei.

**Spectral Density Values**—The values of the spectral density of atomic motions measured at 0, 61, 81, 522, and 696 MHz are plotted in Fig. 5. The high frequency values make a small, relatively uniform contribution to the relaxation across the sequence except at the N terminus, where a significant increase in high frequency motions is observed for the first 10 residues at 522 MHz (Fig. 5*a*). This is in contrast to urea-unfolded apomyoglobin in which minima in  $J(0.87\omega_H)$  plots correspond to maxima in buried surface area in the folded protein (102), suggesting that

hydrophobic interactions persist even in 8 M urea. The lack of definition in the  $J$  (522) and  $J$  (696) plots for reduced Tat-(1-72) suggest a lack of formation of any residual structure at pH 4.1.

The spectral density profiles at 61 and 81 MHz are highly similar and do show some variation with sequence (Fig. 5b). The ends of the protein, residues Lys-61 to Leu-63 and Thr-43, exhibit the smallest contributions at midfrequencies. Interestingly, in the acid-denatured state of apomyoglobin, maxima in buried surface area correlate weakly with maxima in the  $J(\omega_N)$  plot (104), suggesting that  $J(\omega_N)$  is sensitive to formation of transient folding nuclei. The smaller  $J(\omega_N)$  values in Tat near Lys-61 and at the termini suggest more flexibility in these regions. However, there is little evidence of transient folding in the  $J(\omega_N)$  data for Tat-(1-72).

The low frequency spectral densities cover a wider range of values but also contain the highest levels of error in comparison with the values calculated at high frequencies (Fig. 5c). This arises from the uncertainties involved in the  $T_{1\rho}$  measurements. The most notable feature is a peak in slow motions centered at the His<sub>6</sub> affinity tag. There are also less well defined peaks in the proline-rich region and in the basic region. The Cys-rich segment contains the fewest measurements, and they are associated with some of the largest errors. The errors arise from the weak

peak intensities in this region of the protein, probably indicating the presence of conformational exchange in this region.

**Model-free Analysis**—Lipari-Szabo model-free analysis of the relaxation measurements is shown in Fig. 6 and in the supplemental data. Calculations were done both separately for the data collected at each field (supplemental data) and for the two fields combined (Fig. 6) and both with and without  $R_{ex}$  terms. The following observations were made in all analyses. In general, the  $R_{ex}$  contributions were less than  $1\text{ s}^{-1}$  for most residues, the largest value being  $2.5\text{ s}^{-1}$ . The average  $S^2$  values were 0.55 and 0.50 for calculations with and without  $R_{ex}$ , respectively. These results are very similar to analyses of relaxation data in other unfolded proteins, where order parameters in the range of 0.4–0.6 were determined (82, 105). Although the most ordered part of the protein is found in the His<sub>6</sub> affinity tag, some of the highest errors in the analysis are found here. This region also exhibited the highest  $R_2$  values, suggesting that accounting for conformational exchange or hydrogen exchange, not detectable by  $R_{1\rho}$ , may improve the analysis. The  $\tau_c$  values show a slight bell-shaped variation with sequence (supplemental data), the correlation times at the ends of protein being smaller than in the center. Furthermore, the average  $\tau_c$  values (2.1 and 3.3 ns, calculated with and without  $R_{ex}$ , respectively) are barely a factor of 10 greater than the average  $\tau_e$  values (0.19 and 0.25 ns), and in some cases the errors in the internal and overall correlation times overlap. Both of these observations support the notion that HIV-1 Tat-(1-72) exists in a random coil conformation in which there is no clear separation of internal and overall rotational correlation times (97).

**Effect of pH**—The NMR samples used for the assignment and relaxation analysis of Tat were stable for over 1 year. Fig. 7 shows the effects of pH on the  $^1\text{H}/^{15}\text{N}$  HSQC spectrum of Tat-(1-72). The most obvious result is an overall reduction of cross-peak intensities. Fast hydrogen exchange with water might account for this, because exchange at rates greater than  $10^3\text{ s}^{-1}$  will result in the loss of signal intensity from chemical exchange line broadening. To determine if the observed losses are a result of the increasing rates of hydrogen exchange, the theoretical hydrogen exchange rates in unstructured Tat at various pH values were calculated, taking into account the nearest neighbor inductive and steric effects (106) (not shown). For unstructured Tat, the exchange rates between pH 3.3 and 5.8 are

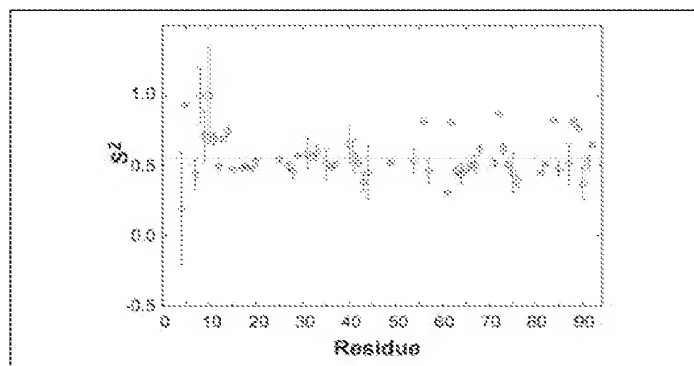


FIGURE 6. Order parameters determined from the two-field Schurr-Lipari-Szabo model-free method with  $R_{ex}(O)$  (mean value, dashed line). All errors in the order parameters were determined based on 100 Monte Carlo simulations of the threshold acceptance method of the simulated annealing optimization.

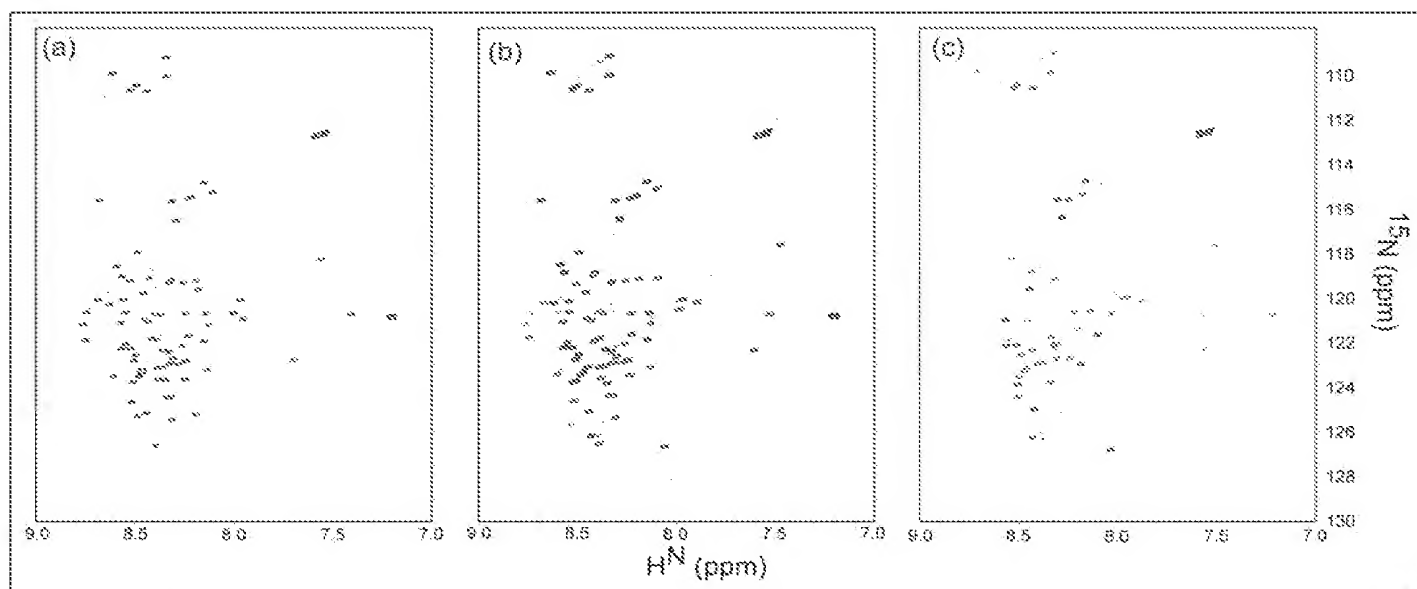


FIGURE 7. Two-dimensional  $^1\text{H}/^{15}\text{N}$  HSQC spectra of Tat-(1-72) at 293 K observed at pH 3.3 (a), pH 4.1 (b), and pH 5.8 (c). All samples are  $\sim 1\text{ mM}$  and were obtained from a single expression/purification. Each spectrum was collected with 32 transients,  $2048 \times 256$  complex points, and sweep widths of 10 ppm in  $F_2(^1\text{H})$  and 24 ppm in  $F_1(^{15}\text{N})$ .



predicted to increase on the order of 250-fold. The calculated exchange rates at pH 5.8 range from  $0.91 \text{ s}^{-1}$  for Gln-92 to  $3.8 \times 10^3 \text{ s}^{-1}$  for Gly-2, and it is likely that for many of the peaks, the intensity loss is attributable to rapid hydrogen exchange. For example, the histidines in the affinity tag and the Cys residues are predicted to be the fastest exchanging amides, and their cross-peaks are lost early in the pH titration. In these cases, the loss of cross-peaks from the spectra is thus indirect evidence that a residue is not involved in a stable, folded conformation. However, detailed analysis of the peaks shows that hydrogen exchange alone cannot explain all of the peak heights (e.g. Thr-40 and Gly-64 are predicted to exchange slowly, yet they disappear early in the titration (Fig. 7)). These results suggest that some cross-peaks lose intensity because of the development of local conformations that are in intermediate exchange on the millisecond to microsecond time scale as observed in the molten globules of two other proteins (58).

**Conclusions**—Natively unfolded proteins have been classified into two categories, depending on whether they exist in a random coil conformation or in a condensed premolten globule (32). Measurement of both static and dynamic multinuclear NMR parameters shows that Tat-(1-72) exists predominantly in a random coil conformation at pH 4.1. However, multinuclear NMR has also uncovered evidence for multiple backbone conformations mainly, but not exclusively, in the Cys-rich region and core. The possible origin of the minor cross-peaks includes *cis-trans* proline peptide isomerization, minor Cys oxidation, and multiple conformers in slow equilibrium. The multiplicity of some peaks in the spectra, together with broadened peaks and the changes in peak intensity as a function of pH suggest that the Cys-rich and core regions form transiently stabilized structures at acidic and neutral pH. The present results are pertinent to Tat interactions with intracellular binding partners, such as cyclin T1, that are expected to encounter only reduced Tat in the intracellular environment. Cyclin T1 probably recognizes the transiently stabilized structure that forms in the Cys-rich region of Tat. Furthermore, the affinity of Tat for the loop region of TAR is greatly increased by interaction with cyclin T1, suggesting binding-induced folding (107), another feature of natively unfolded proteins. Multinuclear NMR is likely to be of value in determining the structures of complexes of Tat and its interaction partners. Finally, there is considerable interest in developing a Tat vaccine (70) based on the presence of Tat antibodies in HIV-1-infected individuals who are long term nonprogressors to AIDS. The antibodies raised against oxidized protein putatively recognize conformational epitopes, suggesting that at neutral pH parts of the protein exist in a stable conformation. The present dynamics analysis suggests that the most likely region to fold is the Cys-rich region, and the formation of disulfide bonds could stabilize local structure there. However, the high positive charge density, paucity of hydrophobic residues, and present dynamics analysis suggest that the remainder of the protein is unlikely to form a stable conformation even at neutral pH.

**Acknowledgments**—We thank Leo Spyropoulos for providing a Mathematica notebook for the single field spectral density and Lipari-Szabo calculations, Walter Englander for providing the Excel files for calculating hydrogen exchange rates, Ryan McKay for helping with the NMR acquisitions on the 800-MHz spectrometer, Richard Sparling for the use of his glove bag, and Vincent Chen for measuring the mass spectra. We are also grateful to Gillian Henry for constructing the expression plasmid and to Avindra Nath for helpful discussions in the early phases of this project.

## REFERENCES

- Lassen, K., Han, Y., Zhou, Y., Siliciano, J., and Siliciano, R. F. (2004) *Trends Mol. Med.* **10**, 525–531
- Arya, S. K., Guo, C., Josephs, S. F., and Wong-Staal, F. (1985) *Science* **229**, 69–73
- Jeang, K. T., Xiao, H., and Rich, E. A. (1999) *J. Biol. Chem.* **274**, 28837–28840
- Berkhout, B., Silverman, R. H., and Jeang, K. T. (1989) *Cell* **59**, 273–282
- Yamaguchi, Y., Takagi, T., Wada, T., Yano, K., Furuya, A., Sugimoto, S., Hasegawa, J., and Handa, H. (1999) *Cell* **97**, 41–51
- Bourgeois, C. F., Kim, Y. K., Churcher, M. J., West, M. J., and Karn, J. (2002) *Mol. Cell Biol.* **22**, 1079–1093
- Kim, Y. K., Bourgeois, C. F., Isel, C., Churcher, M. J., and Karn, J. (2002) *Mol. Cell Biol.* **22**, 4622–4637
- Schulte, A., Czudnochowski, N., Barboric, M., Schonichen, A., Blazek, D., Peterlin, B. M., and Geyer, M. (2005) *J. Biol. Chem.* **280**, 24968–24977
- Bauriwarth, S., and Gagnon, A. (2005) *Curr. HIV Res.* **3**, 61–71
- Mujtaba, S., He, Y., Zeng, L., Farooq, A., Carlson, J. E., Ott, M., Verdin, E., and Zhou, M. M. (2002) *Mol. Cell* **9**, 575–586
- Pumfery, A., Deng, L., Maddukuri, A., de la Fuente, C., Li, H., Wade, J. D., Lambert, P., Kumar, A., and Kashanchi, F. (2003) *Curr. HIV Res.* **1**, 343–362
- Guo, X., Kameoka, M., Wei, X., Roques, B., Gotte, M., Liang, C., and Weinberg, M. A. (2003) *Virology* **307**, 154–163
- Alimonti, J. B., Ball, T. B., and Fowke, K. R. (2003) *J. Gen. Virol.* **84**, 1649–1661
- Gallo, R. C. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 8324–8326
- Pocernich, C. B., Sultana, R., Mohammad-Abdul, H., Nath, A., and Butterfield, D. A. (2005) *Brain Res Brain Res Rev.* **50**, 14–26
- Vendel, A. C., and Lumb, K. J. (2004) *Biochemistry* **43**, 904–908
- Vendel, A. C., and Lumb, K. J. (2003) *Biochemistry* **42**, 910–916
- Neuveut, C., and Jeang, K. T. (1996) *J. Virol.* **70**, 5572–5581
- Derse, D., Carvalho, M., Carroll, R., and Peterlin, B. M. (1991) *J. Virol.* **65**, 7012–7015
- Kuppuswamy, M., Subramanian, T., Srinivasan, A., and Chinnadurai, G. (1989) *Nucleic Acids Res.* **17**, 3551–3561
- Garcia, J. A., Harrich, D., Pearson, L., Mitsuyasu, R., and Gaynor, R. B. (1988) *EMBO J.* **7**, 3143–3147
- Smith, S. M., Pentlicky, S., Klase, Z., Singh, M., Neuveut, C., Lu, C. Y., Reitz, M. S., Jr., Yarchon, R., Marx, P. A., and Jeang, K. T. (2003) *J. Biol. Chem.* **278**, 44816–44825
- Bieniasz, P. D., Grdina, T. A., Bogerd, H. P., and Cullen, B. R. (1998) *EMBO J.* **17**, 7056–7065
- Chen, D., Wang, M., Zhou, S., and Zhou, Q. (2002) *EMBO J.* **21**, 6801–6810
- Weeks, K. M., Ampe, C., Schultz, S. C., Steitz, T. A., and Crothers, D. M. (1996) *Science* **249**, 1281–1285
- Gupta, B., Levchenko, T. S., and Torchilin, V. P. (2005) *Adv. Drug Deliv. Rev.* **57**, 637–651
- Campbell, G. R., Pasquier, E., Watkins, J., Bourgarel-Rey, V., Peyrot, V., Esquieu, D., Barbier, P., de Mareuil, J., Braguer, D., Kaleebu, P., Yirell, D. L., and Loret, E. P. (2004) *J. Biol. Chem.* **279**, 48197–48204
- Avraham, H. K., Jiang, S., Lee, T. H., Prakash, O., and Avraham, S. (2004) *J. Immunol.* **173**, 6228–6233
- Linding, R., Russell, R. B., Neduvu, V., and Gibson, T. J. (2003) *Nucleic Acids Res.* **31**, 3701–3708
- Uversky, V. N., Gillespie, J. R., and Fink, A. L. (2000) *Proteins* **41**, 415–427
- Fink, A. L. (2005) *Curr. Opin. Struct. Biol.* **15**, 35–41
- Uversky, V. N. (2002) *Protein Sci.* **11**, 739–756
- Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F., and Jones, D. T. (2004) *J. Mol. Biol.* **337**, 635–645
- Gunasekaran, K., Tsai, C. J., Kumar, S., Zanuy, D., and Nussinov, R. (2003) *Trends Biochem. Sci.* **28**, 81–85
- Nakayama, K. I., Hatakeyama, S., and Nakayama, K. (2001) *Biochem. Biophys. Res. Commun.* **282**, 853–860
- Bayer, P., Kraft, M., Eichart, A., Westendorp, M., Frank, R., and Rosch, P. (1995) *J. Mol. Biol.* **247**, 529–535
- Gregoire, C., Peloponese, J. M., Jr., Esquieu, D., Opi, S., Campbell, G., Solomiac, M., Lebrun, E., Lebreton, J., and Loret, E. P. (2001) *Biopolymers* **62**, 324–335
- Peloponese, J. M., Jr., Gregoire, C., Opi, S., Esquieu, D., Sturgis, J., Lebrun, E., Meurs, E., Collette, Y., Olive, D., Aubertin, A. M., Witrow, M., Pannecouque, C., De Clercq, E., Bailly, C., Lebreton, J., and Loret, E. P. (2000) *C. R. Acad. Sci. III* **323**, 883–894
- Freund, J., Vertes, L., Koller, K. P., Wolber, V., Heintz, D., and Kalbitzer, H. R. (1995) *J. Mol. Biol.* **250**, 672–688
- Puglisi, J. D., Tan, R., Calnan, B. J., Frankel, A. D., and Williamson, J. R. (1992) *Science* **257**, 76–80
- Long, K. S., and Crothers, D. M. (1999) *Biochemistry* **38**, 10059–10069
- Aboul-ela, F., Karn, J., and Varani, G. (1995) *J. Mol. Biol.* **253**, 313–332
- Seewald, M. J., Metzger, A. U., Willbold, D., Rosch, P., and Sticht, H. (1998) *J. Biomol. Struct. Dyn.* **16**, 683–692
- Metzger, A. U., Bayer, P., Willbold, D., Hoffmann, S., Frank, R. W., Goody, R. S., and Rosch, P. (1997) *Biochem. Biophys. Res. Commun.* **241**, 31–36
- Greenbaum, N. L. (1996) *Structure* **4**, 5–9
- Hakansson, S., and Caffrey, M. (2003) *Biochemistry* **42**, 8999–9006
- Gregoire, C. J., and Loret, E. P. (1996) *J. Biol. Chem.* **271**, 22641–22646

48. Bertoncini, C. W., Jung, Y. S., Fernandez, C. O., Hoyer, W., Griesinger, C., Jovin, T. M., and Zweckstetter, M. (2005) *Proc. Natl. Acad. Sci. U. S. A.* **102**, 1430–1435
49. Tafer, H., Hiller, S., Hilty, C., Fernandez, C., and Wuthrich, K. (2004) *Biochemistry* **43**, 860–869
50. Frank, M. K., Clore, G. M., and Gronenborn, A. M. (1995) *Protein Sci.* **4**, 2605–2615
51. Schwalbe, H., Fiebig, K. M., Buck, M., Jones, J. A., Grimshaw, S. B., Spencer, A., Glaser, S. J., Smith, L. J., and Dobson, C. M. (1997) *Biochemistry* **36**, 8977–8991
52. Dyson, H. J., and Wright, P. E. (1998) *Nat. Struct. Biol.* **5**, 499–503
53. Dyson, H. J., and Wright, P. E. (2001) *Methods Enzymol.* **339**, 258–270
54. Gillespie, J. R., and Shortle, D. (1997) *J. Mol. Biol.* **268**, 170–184
55. Klein-Seetharaman, J., Oikawa, M., Grimshaw, S. B., Wirmmer, J., Duchardt, E., Ueda, T., Imoto, T., Smith, L. J., Dobson, C. M., and Schwalbe, H. (2002) *Science* **295**, 1719–1722
56. Teilum, K., Kragelund, B. B., and Poulsen, F. M. (2002) *J. Mol. Biol.* **324**, 349–357
57. Onuchic, J. N., Wolynes, P. G., Luthey-Schulten, Z., and Socci, N. D. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92**, 3626–3630
58. Redfield, C. (2004) *Methods* **34**, 121–132
59. Dyson, H. J., and Wright, P. E. (2005) *Methods Enzymol.* **394**, 299–321
60. Frankel, A. D., and Pabo, C. O. (1988) *Cell* **55**, 1189–1193
61. Marley, J., Lu, M., and Bracken, C. (2001) *J. Biomol. NMR* **20**, 71–75
62. Kay, L. E., Keifer, P., and Saarinen, T. (1992) *J. Am. Chem. Soc.* **114**, 10663–10665
63. Wittekind, M., and Mueller, L. (1993) *J. Magn. Reson. Ser. B* **101**, 201–205
64. Grzesiek, S., and Bax, A. (1992) *J. Am. Chem. Soc.* **114**, 6291–6293
65. Ikura, M., Kay, L. E., and Bax, A. (1990) *Biochemistry* **29**, 4659–4667
66. Yamazaki, T., Lee, W., Revington, M., Mattiello, D. L., Dahlquist, F. W., Arrowsmith, C. H., and Kay, L. E. (1994) *J. Am. Chem. Soc.* **116**, 6464–6465
67. Vuister, G. W., and Bax, A. (1993) *J. Am. Chem. Soc.* **115**, 7772–7777
68. Cavanagh, J. F., Wayne, J., Palmer, A. G., III, and Skelton, N. J. (1996) *Protein NMR Spectroscopy: Principles and Practice*, Academic Press, Inc., San Diego, CA
69. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR* **6**, 277–293
70. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) *Biochemistry* **33**, 5984–6003
71. Shaka, A. J., Keeler, J., Frenkiel, T., and Freeman, R. (1983) *J. Magn. Reson.* **52**, 335–338
72. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Lyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) *J. Biomol. NMR* **6**, 135–140
73. Schwarzing, S., Kroon, G. J., Foss, T. R., Chung, J., Wright, P. E., and Dyson, H. J. (2001) *J. Am. Chem. Soc.* **123**, 2970–2978
74. Smith, L. J., Bolin, K. A., Schwalbe, H., MacArthur, M. W., Thornton, J. M., and Dobson, C. M. (1996) *J. Mol. Biol.* **255**, 494–506
75. Kay, L. E., Nicholson, L. K., Delaglio, F., Bax, A., and Torchia, D. A. (1992) *J. Magn. Reson.* **97**, 359–375
76. Habazettl, J., Myers, L. C., Yuan, F., Verdine, G. L., and Wagner, G. (1996) *Biochemistry* **35**, 9335–9348
77. Opi, S., Peloponese, J. M., Jr., Esqueit, D., Watkins, J., Campbell, G., De Mareuil, J., Jeang, K. T., Yirrell, D. L., Kaleebu, P., and Loret, E. P. (2004) *Vaccine* **22**, 3105–3111
78. Abragam, A. (1961) *The Principles of Nuclear Magnetism*, Clarendon Press, Oxford, UK
79. Farrow, N. A., Zhang, O., Szabo, A., Torchia, D. A., and Kay, L. E. (1995) *J. Biomol. NMR* **6**, 153–162
80. Palmer, A. G. (2004) *Chem. Rev.* **104**, 3623–3640
81. Peng, J. W., and Wagner, G. (1992) *Biochemistry* **31**, 8571–8586
82. Farrow, N. A., Zhang, O., Forman-Kay, J. D., and Kay, L. E. (1997) *Biochemistry* **36**, 2390–2402
83. Szyrski, T., Luginbuhl, P., Otting, G., Guntert, P., and Wuthrich, K. (1993) *J. Biomol. NMR* **3**, 151–164
84. Lefevre, J. F., Dayie, K. T., Peng, J. W., and Wagner, G. (1996) *Biochemistry* **35**, 2674–2686
85. Szyrski, T., Luginbuhl, P., Yee, A. A., and O'Neil, J. D. (1996) *J. Biomol. NMR* **7**, 283–294
86. Wolfram Research, Inc. (2004) *Mathematica*, Version 5.1, Champaign, IL
87. Creighton, T. E. (1993) *Proteins: Structures and Molecular Properties*, 2nd Ed., p. 4, W. H. Freeman, New York
88. Lipari, G., and Szabo, A. (1982) *J. Am. Chem. Soc.* **104**, 4546–4559
89. Andrec, M., Montelione, G. T., and Levy, R. M. (1999) *J. Magn. Reson.* **139**, 408–421
90. Schurr, J. M., Babcock, H. P., and Fujimoto, B. S. (1994) *J. Magn. Reson. Ser. B* **105**, 211–224
91. Yao, J., Dyson, H. J., and Wright, P. E. (1997) *FEBS Lett.* **419**, 285–289
92. Zhang, O., Forman-Kay, J. D., Shortle, D., and Kay, L. E. (1997) *J. Biomol. NMR* **9**, 181–200
93. Peti, W., Smith, L. J., Redfield, C., and Schwalbe, H. (2001) *J. Biomol. NMR* **19**, 153–165
94. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) *J. Mol. Biol.* **222**, 311–333
95. Wishart, D. S., and Sykes, B. D. (1994) *J. Biomol. NMR* **4**, 171–180
96. Zhang, O., and Forman-Kay, J. D. (1995) *Biochemistry* **34**, 6784–6794
97. Hu, Y., Macinnis, J. M., Cherayil, B. J., Fleming, G. R., Freed, K. F., and Perico, A. (1990) *J. Chem. Phys.* **93**, 822–836
98. Ulrich, D. L., Kojetin, D., Bassler, B. L., Cavanagh, J., and Loria, J. P. (2005) *J. Mol. Biol.* **347**, 297–307
99. Thormann, T., Soroka, V., Mielbo, S., Berezin, V., Bock, E., and Poulsen, F. M. (2004) *Biochemistry* **43**, 10364–10369
100. Bhavesh, N. S., Sinha, R., Mohan, P. M., and Hosur, R. V. (2003) *J. Biol. Chem.* **278**, 19980–19985
101. Otzen, D. E., Miton, S., Akke, M., and Oliveberg, M. (2004) *Biochemistry* **43**, 12964–12978
102. Schwarzing, S., Wright, P. E., and Dyson, H. J. (2002) *Biochemistry* **41**, 12681–12686
103. Platt, G. W., McParland, V. J., Kalverda, A. P., Homans, S. W., and Radford, S. E. (2005) *J. Mol. Biol.* **346**, 279–294
104. Yao, J., Chung, J., Eliezer, D., Wright, P. E., and Dyson, H. J. (2001) *Biochemistry* **40**, 3561–3571
105. Buevich, A. V., Shinde, U. P., Inouye, M., and Baum, J. (2001) *J. Biomol. NMR* **20**, 233–249
106. Bai, Y., Milne, J. S., Mayne, L., and Englander, S. W. (1993) *Proteins* **17**, 75–86
107. Wei, P., Garber, M. E., Fang, S. M., Fischer, W. H., and Jones, K. A. (1998) *Cell* **92**, 451–462